DOCUMENT RESUME

ED 360 828 FL 021 376

AUTHOR	Brecht, Richard D.; And Others
TITLE	Predictors of Foreign Language Gain during Study Abroad. NFLC Occasional Papers.
INSTITUTION	Johns Hopkins Univ., Washington, DC. National Foreign Language Center.
PUB DATE	93
NOTE	33p.
AVAILABLE FROM	National Foreign Language Center, 1619 Massachusetts Ave., N.W., Washington, DC 20036 ($5 prepaid).
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Achievement Gains; College Students; Graduate Students; Grammar; Higher Education; *Language Proficiency; *Language Skills; Predictor Variables; Reading Skills; Second Language Learning; *Second Languages; Sex Differences; Student Characteristics; *Study Abroad

ABSTRACT

A large-scale statistical study of the predictors of language gain during study abroad is reported. Subjects were 658 American college and graduate students in the Soviet Union between spring 1984 and spring 1990. Variables examined included student characteristics (age, gender, citizenship, country of birth, place and levels of formal education, highest degree taken, major, prior Russian and other second language training and experience, program type, overseas host institution) and a variety of language measures (proficiency tests, program qualifying exams, learning style or aptitude data). Statistical procedures used are outlined, and results are charted and discussed. It was found that certain student characteristics were predictive of language gains abroad, including gender, experience in learning other foreign languages, and command of grammar and reading skills. Implications are seen for student preparation and selection, in-country program design, and second-language instruction in general. A brief bibliography and author notes are included. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 360 828

Predictors of Foreign Language Gain during Study Abroad

RICHARD D. BRECHT,
DAN DAVIDSON,
and RALPH B. GINSBERG
The National Foreign Center
and
The American Council of Teachers of Russian

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R D Lambert

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

FL021376

The National Foreign Language Center
Washington DC

BEST COPY AVAILABLE

About the Occasional Papers

This is one in a series of Occasional Papers published by the National Foreign Language Center. The NFLC prints and distributes articles on a wide variety of topics related to foreign language research, education, and policy. The Occasional Papers are intended to serve as a vehicle of communication and a stimulant for discussion among groups concerned with these topic areas. The views expressed in these papers are the responsibility of the authors and do not necessarily reflect the views of the NFLC or of the Johns Hopkins University.

About the National Foreign Language Center

The National Foreign Language Center, a nonprofit organization established within the Johns Hopkins University in 1987 with support from major private foundations, is dedicated to improving the foreign language competency of Americans. The NFLC emphasizes the formulation of public policy to make our language teaching systems responsive to national needs. Its primary tools in carrying out this objective are:

— *Surveys.* The NFLC conducts surveys to collect previously unavailable information on issues concerning national strength and productivity in foreign language instruction, and our foreign language needs in the service of the economic, diplomatic, and security interests of the nation.

— *National policy planning groups.* In order to address major foreign language policy issues, the NFLC convenes national planning groups that bring together users of foreign language services and representatives of the language instructional delivery systems in formal education, the government, and the for-profit sector.

— *Research.* The NFLC conducts research on innovative, primarily individual-oriented strategies of language learning to meet the nation's foreign language needs of the future.

In addition, the NFLC maintains an Institute of Advanced Studies where individual scholars work on projects of their own choosing.

The results of these surveys, discussions, and research are made available through the NFLC's publications, such as these Occasional Papers, and they form the basis of fresh policy recommendations addressed to national leaders and decision makers.

Copyright © 1993, The National Foreign Language Center at the Johns Hopkins University

Price per copy: $5.00, prepaid

Predictors of Foreign Language Gain during Study Abroad

Richard D. Brecht, Dan Davidson, and Ralph B. Ginsberg

Students, teachers, and policymakers alike assume that truly functional competence in a language requires spending time living in the country where that language is spoken. Whatever else our academic programs can accomplish, the logic goes, classroom drills cannot substitute for extended experience communicating with native speakers in natural settings about real-life matters. This general impression is reinforced by students returning from abroad, who frequently demonstrate significantly improved language skills and who testify to the value of the experience. Indeed, college programs in languages known to be less accessible to English speakers, like Arabic, Chinese, Japanese, and Russian, essentially have come to rely upon in-country language programs to insure even the basic minimum level of functional competence. And French, German, and Spanish too, as a practical matter, assume this in-country stay for their serious students, particularly in light of the fact that travel to Western Europe and Latin America is readily available and relatively affordable.

Until recently, however, to assess the impact of study abroad, students and teachers have had to rely exclusively upon these intuitions and subjective observations.[1] The foreign language acquisition field has been weak in responding to demands for rigorous answers to questions like, Do students actually gain

The research contained in this report was funded by a grant from the U.S. Department of Education to the American Council of Teachers of Russian and the National Foreign Language Center. Further research has continued under a grant from the Ford Foundation to the National Foreign Language Center. Other aspects of this study have been reported in Brecht, Davidson, and Ginsberg 1990, 1991; and Brecht and Davidson 1992. An analysis of the data can be found in Ginsberg 1992 (available from the National Foreign Language Center), which contains a comprehensive summary of the data and analysis procedures reported on here, and will be referred to here in order to allow a more focused presentation of the salient points of the study. The authors wish to acknowledge their debt to Paul Wheeling for assistance in developing and managing the ACTR database.

significantly in language skills from in-country study, particularly in comparison with intensive domestic programs? If they do, which students are most likely to gain? Which are most likely to have difficulty? What precisely is gained linguistically from living in the country? Which language skills? To what levels of proficiency? What kinds of experience and programs in-country are the most effective in building students' language skills? What is the minimal duration an effective program must have? How much language must a student have in order to take maximum advantage of the experience? If resources dictate that most students will have only one chance at this opportunity, what is the most effective time in one's language-learning career to spend in-country? How, in fact, does language learning in-country differ substantially from other forms of language acquisition?

1. INTRODUCTION

This paper presents the first large-scale statistical study of the predictors of language gain in the study-abroad environment. It shows that certain characteristics of students going abroad are significant predictors of successful language learning, among them gender, experience in learning other foreign languages, and strong command of grammar and reading skills. Clearly, such results have major policy and scientific import. From the policy point of view, such results have significant ramifications for student preparation and selection as well as for in-country program design. On the scientific side, such results speak to basic hypotheses in the field that heretofore have been addressed on a rather intuitive basis, or at best on the basis of very small samples with little rigorous measurement. More important, the issues that are addressed by these data have application for second language acquisition in general, given the fact that the study-abroad programs under investigation focus on advanced-level skills in Russian—a language of "Class 3" difficulty—and consist of a structured academic program as well as substantial exposure to native speakers in natural settings. This blend of structured and unstructured learning is increasingly recognized as representative of the typically successful language-learning career of foreign language students, which necessarily relies on a combination of academic programs, individual study, and intensive interaction with native speakers of the language.[2]

The present paper is part of a series based on collaborative research sponsored by the American Council of Teachers of Russian (ACTR) and the National Foreign Language Center addressing these and similar questions. It presents the results of a systematic analysis of an extensive database on university students studying advanced Russian for one semester in eight educational institutes in Moscow and St. Petersburg (formerly Leningrad) assembled by ACTR over a period of more than fifteen years.[3] This database is unique in its ability to provide evidence for understanding language acquisition in study abroad:

— It contains carefully collected before-and-after measures of three of the four basic language modalities (speaking, listening, and reading), so that gains can be assessed.
— It is rich in personal and learning variables—including personal data provided on the student application forms, grades of courses at the various institutes, and reports on student performance by ACTR academic program officers in the former Soviet Union—so that determinants and correlates of gain can be assessed.[4]
— It is extensive, so that conclusions are defensible—that is, not based on small numbers—with considerable statistical control exercised.[5]
— It has been systematically converted to machine-readable form and entered into a database management system.

This paper is meant to document the data used, the analytical strategy and methods employed, and the most salient conclusions that follow from this unprecedented empirical study, as well as the grounds on which substantive conclusions from the data rest. Moreover, this first long-term, broadly cross-sectional, empirical study of adult in-country second language acquisition can be linked directly to data collected in the Ford Foundation–sponsored ethnographic study of language acquisition in study abroad conducted by the National Foreign Language Center in cooperation with ACTR. This connection makes it possible to relate what students bring to the program to information on in-country language-learning behavior, and hence to understand better the conditions under which improvements in language competence occur.

In the following sections we select for discussion the most salient variables having an important effect on language learning: gender, knowledge of other foreign languages, general language-learning aptitude as measured by the Modern Language Aptitude Test (MLAT), and prior proficiency in other language modalities, specifically grammar and reading. Many other variables were included in the analysis and will be mentioned, when important, and displayed in the summary tables. In all cases, discussion of the role of any one variable implies that its effect is measured after all other variables are held constant.

2. SUBJECTS, VARIABLES, AND METHODS

2.1. Subjects

Since 1976 ACTR has maintained records pertaining to the general academic and in-country language performance of more than two thousand American undergraduate and graduate students who, under its auspices, have completed long-term language training programs in the former Soviet Union. During most of the recording period participants have typically been at the B.A. or immediate post-B.A. level, 22.2 years of age, with an undergraduate major or equivalent preparation in Russian language and area studies, with or without other aca-

demic specializations. Because of limitations imposed by the previous government on the number of Americans permitted to study in Russia, competition in the United States for places in the ACTR programs was keen, with as many as four to six qualified applicants—often Russian majors or graduate students in the field—for each position abroad.[6]

A steady increase in the number of positions for study in Russia since 1985—from 110 in 1984–85 to 480 in 1989–90—has been accompanied by a doubling in the number of sending institutions in the United States, reflecting a greater diversity of institutional types and geography than was the case in earlier years. At the present time a total of 195 colleges and universities have placed students or faculty in the ACTR programs. Represented in this number are geographically diverse public and private institutions, including small colleges and large comprehensive universities. Among those institutions ranked highest for the number of students sent on ACTR long-term study programs, there is a diversity of types of Russian language programs to be found. Of the sixty-five highest-ranking institutions, no single one accounts for more than 6 percent of the total ACTR database.

The analysis in this report is based on data relating to 658 students who studied in four-month ACTR programs in Russia (the former Soviet Union) from the spring of 1984, when ACTR started administering oral proficiency interviews (OPIs) before and after the program, through the spring semester of 1990.[7] Since the students are essentially self-selected within the standards set by ACTR, the conclusions here cannot be generalized with confidence to randomly selected students of Russian who might have the opportunity to study abroad. Nevertheless, to the extent that ACTR's selection criteria are controlled in the analyses—which, for the most part, they are in the regression analyses of gain—the results hold more generally.[8]

The specific data selected for analysis in the present study include data on age, gender, citizenship, country of birth, place and levels of formal education (including highest degree taken), major field(s), and information on all prior training in the Russian language, including number of contact hours, prior experience abroad, experience in intensive stateside courses, secondary school programs, language laboratory use, and knowledge of Slavic and non-Slavic languages in addition to English and Russian. Also included in the database for each student record is the academic year and type of program completed and the host institution in Russia.

2.2. *Language Measures*

The instruments measuring language proficiency in various modalities are described thoroughly in Brecht, Davidson, and Ginsberg (1990, 1991). The OPI and proficiency-oriented tests of listening and reading developed by the Educational Testing Service—ETS Listening and ETS Reading (ETSL and ETSR)—were ad-

ministered just before and at the very end of the program and are the basis of measures of gains.[9] The ACTR qualifying exams, measuring achievement in grammar and reading, supplement the proficiency measures determining pre-program levels. Learning-style ("aptitude") data are included in this study in the form of both raw scores and three subtest scores of the Short Form MLAT (MLATSF)—ability to use analytic (MLAT3), synthetic (MLAT4), and memory-based (MLAT5) learning strategies—which are analyzed separately to see exactly which aptitudes and strategy configurations might affect gains.

Over the period analyzed here the language data collected by ACTR varied somewhat. As a consequence, the analyses are based on different numbers of

Table 1
Administration, Missing Data, and Valid Cases for Various Instruments, by Program Date

		Number Missing				
Program	Total Obs.	Change OPI	Change ETSL	Change ETSR	ACTR Qual. †	MLAT
Spring 1984	19	0	*	*	1	*
Fall 1984	18	0	*	*	2	1
Spring 1985	21	0	*	*	5	0
Fall 1985	13	0	*	*	0	0
Spring 1986	19	0	0	0	0	0
Fall 1986	20	0	3	2	1	0
Spring 1987	42	1	0	1	2	4
Fall 1987	36	1	1	1	0	2
Spring 1988	49	2	1	1	4	0
Fall 1988	100	2	2	2	4	4
Spring 1989	100	1	2	2	12	3
Fall 1989	119	2	6	6	23	4
Spring 1990	102	3	9	9	6	*
Valid cases	658	646	563	563	598	519

* Indicates that the instrument was not administered.

† The numbers refer to the composite score QualGen. For analyses in which the qualifying grammar and reading tests are used as separate variables (thus using students before spring 1990 only), numbers in the table give missing data for one or the other or both, but the base numbers are essentially the same.

cases, depending on the variables involved. ACTR started administering the OPI in the spring of 1984. The ETS exams were phased in for spring 1986. The MLATs began in the fall of 1984 and were discontinued in the spring of 1990. In spring 1990 the qualifying reading and grammar tests were combined into one test. For students before the spring of 1990, scores on the qualifying grammar and reading tests have been combined into a composite score, comparable to that of spring 1990, referred to below as QualGen.[10] Moreover, there is a small amount of missing data on all of the instruments. For example, it is not possible to calculate change scores for twelve students since they did not take the OPI, but data on these students can be used to study change in ETSL and ETSR. Table 1 presents a summary of what language data are available, by program date.

Descriptive statistics for the quantitative preprogram language measures are given in Table 2. Four fifty-minute tests (two forms each for reading and listening) are designed to be reliable as measures of reading and listening comprehension in the intermediate/high to superior range (1+ to 3 levels). Stimulus material in both sets of tests is drawn from a variety of natural language use sources, relying in particular on materials taken from the mass media. Reading passages are printed texts in Russian, ranging from short passages designed to assess extraction of factual information to larger texts designed to measure comprehension and analysis. Listening passages, administered by means of a tape recording, contain material in Russian, spoken by both males and females at normal speed, such as news broadcasts, interviews, and conversations.

OPI speech samples have been recorded on audio- or videotape for subsequent verification and analysis. "Plus" ratings (0+, 1+, 2+, 3+, etc.) are given to

Table 2
Descriptive Statistics for Preprogram Language Measures

Variable	Median	Mean	Standard Deviation	Range
ETSL (pre)	21	22.1	8.8	1–49
ETSR (pre)	19	19.7	8.1	2–45
Qualifying grammar	642	636.2	141.3	156–987
Qualifying reading	655	646.5	177.7	88–985
Qualifying general	640	630.0	144.6	187–975
MLAT3	32	32.8	8.9	6–50
MLAT4	32	31.2	5.6	10–44
MLAT5	23	21.1	4.1	4–24
MLATSF (total)	86	85.2	13.0	29–115

Table 3
Frequency Distributions for Preprogram Ratings on OPI and ETS Listening and Reading

Test	0	0+	1	1+	2	2+	3	Total
OPI	4	49	378	134	60	21	4	660
%	0.1	7.4	57.3	20.3	9.1	3.2	0.1	100
ETSL	—	—	482	63	18	7	8	578
%	—	—	83.4	10.9	3.1	1.2	1.4	100
ETSR	—	—	116	257	111	41	52	577
%	—	—	20.1	44.5	19.2	7.1	9.2	100

students who substantially surpass the requirements for a given level but fail to sustain performance at the next higher level; "mid"-level ratings (novice-mid, intermediate-mid) are recorded in the ACTR database simply as 0 (novice) or 1 (intermediate). All OPI test results in the ACTR data are based on face-to-face interviews conducted by U.S. testers certified by the American Council on the Teaching of Foreign Languages (ACTFL) who are not otherwise involved in the in-country training programs. The incidence of unratable samples has been relatively low.

Frequency distributions for preprogram OPI and the proficiency-coded ETS tests are given in Table 3. For purposes of comparison, in the table the two ETS tests have been normed by ETS to make them comparable to the levels of the OPI. Note that both ETS tests have a ceiling at 3—that is, students cannot get a rating above this level—and a floor at 1. (On the OPI no one happened to score above 3.) The analysis of reading and listening test results, however, focuses on the raw scores, which are well spread and unbounded.

2.3. *Learner/Student Characteristics in the ACTR Database*

Table 4 presents descriptive statistics for the personal characteristics and educational history variables that are used in the explanation of gains. (The host institution in Russia is also included.) Three variables—country of birth, number of class hours, and number of lab hours—were examined but are not used here because in preliminary studies they added no explanatory power to the other variables in the table. "Resident director ratings" are included in the database, but they were not included in the present paper.[11] For many of the variables there

is a small amount of missing data; these values were imputed, as indicated in the table, so as to retain the cases in the analysis.[12]

2.4. *Analytical Methods*

The methods of analysis used in this study follow standard statistical practice. Where prediction and explanation are concerned—as in the analysis of the determinants of gain—regression analysis and discriminant analysis are used. For the most part, regressions are estimated by ordinary least squares (OLS), but in the analysis of OPI gains (Section 3.5), binary and ordinal probit regressions are also used.[13] In the regression analyses, conclusions are based on models that control for all available relevant factors—that is, the relationships hold over and

Table 4
Descriptive Statistics for Student Characteristics

Variable	Summary Statistics	Number Missing	Treatment of Missing Data
Age	median 21; 75% between 20 and 22; range 17–33	0	—
Gender	58% women	0	—
Country of birth	explored in preliminary runs but not used here	0	—
Undergrad. college	explored as a background factor	3	—
Major	Russian (59%); area studies (16%); humanities (12%); other (13%)	12	imputed as Russian
Degree	40% still undergraduate; 49% B.A.; 11% M.A. or Ph.D.	14	imputed as B.A.
High school Russian	75% none; a scattering above; recoded to 0/1	15	imputed as 0
College Russian	mode = 3: for frequencies see tables below	15	imputed as 3
Class hours	explored in preliminary runs but not used here	15	—
Lab hours	explored in preliminary runs but not used here	15	—
Slavic languages	0/1 variable; 5% know one	19	imputed as 0
Non-Slavic languages	0 (11%); 1 (46%); 2 (31%); 3+ (12%)	13	imputed as 1
Previous immersion	0/1 variable; 25% have a previous immersion	15	imputed as 0
Institute	programs were held at 8 institutes	—	—

above what can be explained by other relevant variables. For each analysis, results have been carefully checked to see that they are not influenced by outliers or leverage points.

In several of the analyses, data have been examined from many points of view to see that the results are stable—that is, not overly dependent on a specific model or method of analysis. This is particularly relevant to analyses involving the OPI, an ordinal variable that does not accommodate gracefully to standard fully quantitative or qualitative statistical models. In presenting the analyses, a blend of tabular and quantitative summaries have been used to make the main conclusions accessible to readers who are not technically trained in data analysis, and at the same time to tell readers who are interested in the statistical results what they need to know.[14]

3. PREDICTORS OF GAIN

3.1. Measures of Gain

For the quantitative variables measuring listening and reading skills (ETSL and ETSR), gain is defined as the difference between preprogram and postprogram scores, and the assessment of factors affecting gain can be accomplished with OLS regression.[15] On both listening and reading proficiency, gain is very strongly, negatively related to preprogram level, with correlations of −.522 for listening and −.344 for reading; that is, the higher the initial level, the less the gain. This phenomenon is consistent with a normal S-shaped learning curve, since the subjects in the present study are beyond initial levels. As a consequence of these strong relationships, analysis of the effects of other variables, such as gender, previous immersion, or MLATs, makes sense only with preprogram levels controlled—that is, in terms of what is not explained by preprogram level.

For the OPI, the situation is considerably more complicated because of the nature of the scale. The OPI score is an ordinal variable; each level should be thought of as a grouping of scores on an underlying unobserved scale of proficiency on which variation is more continuous. Grouping loses information, in effect introducing a measurement error, in that students with quite different (unobserved) proficiency levels could be given the same OPI score. In general, grouping attenuates relationships.[16] Gain on the OPI (the difference between scores before and after), also ordinal, has the same difficulties, difficulties that it inherits from its components. Table 5 shows the relation between pre- and postprogram OPI scores. The main diagonal—the cells enclosed in boxes—represents no gain; moving to the left of the main diagonal in any row indicates a loss; and moving to the right one column (or two, or three) indicates a gain of half a point (or one, or one and a half). As with listening and reading, and for many of the same reasons, the association is strongly negative, with the probability of gaining reduced sharply as initial level increases.

Table 5
OPI Scores, Pre- and Postprogram (Count/Row Percent)

	Post OPI							
Pre OPI	**0+**	**1**	**1+**	**2**	**2+**	**3**	**3+**	**Total**
0	1	3	0	0	0	0	0	4
%	25.0	75.0	0	0	0	0	0	100
0+	1	30	16	1	1	0	0	49
%	2.04	61.2	32.7	2.04	2.04	0	0	100
1	0	83	203	71	16	1	0	374
%	0	22.2	54.3	19.0	4.28	0.27	0	100
1+	0	8	48	48	29	1	0	134
%	0	5.97	35.8	35.8	21.6	0.75	0	100
2	0	0	7	33	18	2	0	60
%	0	0	11.7	55.0	30.0	3.33	0	100
2+	0	0	0	5	6	10	0	21
%	0	0	0	23.8	28.6	47.6	0	100
3	0	0	0	0	0	3	1	4
%	0	0	0	0	0	75.0	25.0	100
Total	2	124	274	158	70	17	1	646
%	0.310	19.2	42.4	24.5	10.8	2.63	0.15	100

In our view, because the number of levels one could gain is so limited, especially at higher initial levels, there is no fully satisfactory way of quantifying gain using the OPI. Instead, analyses with a number of different definitions and procedures for controlling for initial levels were run to get a handle on the effects: results that are not sensitive to the definition and procedure are likely to be significant, whereas results that hold for only one procedure may well be artifacts. Two gain criteria are used in the tables of Section 3.5: a simple no-increase versus increase (coded 0/1) and a three-level variable, the levels being no gain or loss, a gain of one level, and a gain of two or more levels (coded 0/1/2).[17] Of course, for both definitions of gain, preprogram level must be controlled in the analysis.

3.2. *Regression Strategy*

The following three sections present the results of analyses of factors affecting gains on each of the three criteria of language proficiency defined in the ACTR database. The regression strategy in each section is the same. We start with a baseline of student background characteristics, as listed in Table 4, and the preprogram level of the criterion in question. Possible effects of undergraduate college, program date, host institution, highest degree, and undergraduate major

are explored by examining their relation to residuals from the baseline. With the notable exceptions of undergraduate major and program date for oral proficiency, these variables turn out to be generally nonsignificant. The language measures—MLATs, qualifying exam, and preprogram scores on other criteria—are then added to the baseline, separately and in combination, to gauge their effects. The different sets of regressions are based on different numbers of observations because of the missing instrument pattern specified in Table 1, so R^2 across sets cannot be directly compared. Once the factors that affect the gain are identified, nonsignificant variables are eliminated to produce a "good" model in which the coefficients and their standard errors—and hence the t-statistics—are estimated with the greatest possible precision. For reasons of space, only the final, good model is presented here.[18]

3.3. *Predictors of Gain: Listening Proficiency*

Table 6 contains the results of the analysis of factors related to gains in listening proficiency. It presents a good model for the predictors of gains in listening proficiency, with the nonsignificant variables and the highly significant predictors included. The high R^2 indicates that this is an excellent model indeed.

Table 6
Model for Gain in Listening Proficiency with Only Significant Variables Included

	Model 10	
Variable	**Coefficient**	**t-Statistic**
Constant	22.1062	7.61
ETSL1	-0.69437	-16.50
Gender	1.13240	2.04
Age	-0.48846	-4.25
High school Russian	-1.10406	-1.68
Slavic languages	-2.04223	-1.60
Previous immersion	1.40194	2.23
ETSR1	0.36425	7.80
MLAT3	0.14506	4.45

R^2 = .429, adjusted R^2 = .418, d.f. = 444.

t-statistics > 1.65 are significant at the .05 level (one tailed); t's > 1.96 are significant at the .025 level; t's > 2.33 are significant at the .01 level; and t's > 3 are very highly significant ($\alpha > .001$).

As noted above, the preprogram listening proficiency level is a strong determinant of gain, no matter what is controlled, and must be included in any analysis. Over and above what can be accounted for by preprogram level and all of the other variables in the equation, several individual characteristics that relate to the hypotheses (questions) put forward in Section 1 have significant effects, judging from the t-statistics:

— *Gender:* Men gain more than women.
— *Age:* Younger people gain more than older.
— *High school Russian:* People who have had Russian in high school gain less than people who have not.
— *Non-Slavic languages:* People who know other foreign languages gain more than people who do not.
— *Previous in-country immersion:* Students with a previous in-country immersion experience gain more than those without.
— *Preprogram reading proficiency level:* This level is strongly related to gains in listening.
— *MLAT3:* The MLAT3 score is highly significant.

The implications of the gender effect are discussed below in Section 4.2; knowledge of other languages and previous immersion are considered in Section 4.3. (Age is left for future comment.) The lack of an effect of college Russian is probably the result of controlling for preprogram levels, with which it is correlated. (Accordingly, this variable is omitted from Table 6.) The negative effect of having studied Russian in high school is probably an artifact of its interrelations with the other variables in the equation. From Table 6 it is clear that MLAT3 (use of analytic strategies) is highly significant. (A separate test shows that the total MLAT score, MLATSF, cannot substitute for its individual components.) Knowledge of another Slavic language is rare in the sample and is included here only for exploratory purposes.

3.4. *Predictors of Gain: Reading Proficiency*

Two good models for gains in reading proficiency, with nonsignificant baseline variables dropped, are presented in Table 7. (There are two models here because of the different sample sizes when the qualifying exams are included.) The good models are lacking in predictors, with only the MLATS and qualifying reading (QualRead) highly significant. Preprogram listening proficiency is barely significant, and gender is still nonsignificant. Of course, preprogram ETS Reading proficiency is very strongly negatively related to gains, with the ceiling no doubt playing a role. Also, it is possible that the QualRead effect is picking up the difference between true reading proficiency and the preprogram test score, so that the QualRead effect has little substantive meaning. It bears reiteration that

Table 7
Models for Gain in Reading Proficiency with Only Significant Variables Included

	Model 10		Model 11	
Variable	**Coefficient**	**t-Statistic**	**Coefficient**	**t-Statistic**
Constant	-3.39018	-1.65	-5.51360	-2.47
ETSR1	-0.41367	-8.85	-0.48236	-8.98
Gender	0.85530	1.54	0.78725	1.37
ETSL1	0.10968	2.61	0.07314	1.66
% MLATSF	0.15342	6.91	0.15581	6.73
% QualRead	—	—	0.00684	3.40

Model 10: R^2 = .205, adjusted R^2 = .198, d.f.= 448.
Model 11: R^2 = .225, adjusted R^2 = .215, d.f. = 408.

the formal study programs reviewed here place greater emphasis on oral skills than on reading.

3.5. *Predictors of Gain: Oral Proficiency*

For the analysis of gains in oral proficiency, as measured by the OPI, a somewhat different regression strategy from that of the previous two sections was employed. First, as discussed above, the ordinal, but nevertheless qualitative, nature of the scale requires examination of several criteria of gain. Two criteria as defined above are used: 0/1 (no gain/gain) and 0/1/2 (no gain/gain of one level/gain of two or more levels). Second, different regression models are required for the different criteria: probit for 0/1 and ordinal probit for 0/1/2. OLS results are also presented for comparability with previous sections. A given variable may not necessarily have the same effect on the two criteria. As a general rule, a variable that discriminates only between gains of 1 and 2 will be significant for 0/1/2 but nonsignificant for 0/1. By contrast, a variable that discriminates between 0 and 2 will be significant for 0/1 but not for 0/1/2.

Tables 8 and 9 present good models—that is, with nonsignificant variables on both criteria dropped—for 0/1 and 0/1/2, respectively. The goodness of fit of the models is significant but hardly dramatic.[19]

As noted in Section 3.1, a vexing problem in the analysis of OPI gains is how to control adequately for initial levels. A direct method of control is to hold the

Table 8
Model for Gain in Oral Proficiency with Only Significant Variables Included (0/1 Criterion)

	OLS		Probit	
Variable	**Coefficient**	**t-Statistic**	**Coefficient**	**t-Statistic**
Constant	127.934	3.46	488.006	3.56
OPI1	-0.25861	-10.31	-0.86840	-9.04
High school Russian	0.09324	2.18	0.32370	2.03
Non-Slavic languages	0.02616	1.19	0.08816	1.10
ETSL1	0.00235	0.82	0.00873	0.86
ETSR1	0.00763	2.38	0.02924	2.48
Major	0.01324	0.79	0.05490	0.90
Program date	-0.06378	-3.43	-0.24452	-3.55
QualGen	0.00031	1.84	0.00122	2.00

R^2 (OLS) = .231, d.f. = 508.

Log likelihood (probit) = -248.633.

Table 9
Model for Gain in Oral Proficiency with Only Significant Variables Included (0/1/2 Criterion)

	OLS		Ordinal Probit	
Variable	**Coefficient**	**t-Statistic**	**Coefficient**	**t-Statistic**
Constant	226.3520	3.95	424.0150	3.94
OPI1	-0.40553	-10.44	-0.76746	-9.72
High school Russian	0.14042	2.12	0.273361	2.23
Non-Slavic languages	0.05106	1.49	0.09329	1.48
ETSL1	0.00738	1.66	0.01311	1.59
ETSR1	0.01184	2.39	0.02237	2.40
Major	0.05234	2.01	0.10336	2.13
Program date	-0.11265	-3.91	-0.21249	-3.93
QualGen	0.00056	2.17	0.00104	2.16

R^2 (OLS) = .251, d.f. = 508.

Log likelihood (probit) = -465.656.

μ = 1.566 (t = 18.69).

initial level constant by running the analysis separately for different initial groups.[20] Students with preprogram OPIs of 1 and 1+ are numerous enough to analyze. Results for initial OPIs of 1 are generally consistent with those of the whole sample—which is not surprising, in that they constitute the large majority of students. On the 0/1 criterion in the expanded baseline, only program date is significant. The Slavic language variable is borderline but negative ($t = -1.6$) and is probably not worth interpreting. ETS Reading and QualGen are highly significant, but ETS Listening is not significant. With the 0/1/2 criterion, major and ETS Listening are significant, as in the whole sample.

The situation is more interesting with people who start at 1+. Only the 0/1 criterion, which represents a significant gain to 2 ("advanced") or above, is used because of the small sample size. Table 10 gives the results from the OLS regression. Gender is highly significant, with men more likely than women to reach 2: this accounts for the incipient gender effects in the whole sample and has important implications (see Section 4.2 below for discussion). Both knowledge of Slavic and knowledge of other foreign (non-Slavic) languages have positive effects, indicating the possible effect of previous language-learning experience (see Section 4.3 below). QualGen is very significant, while MLATs and proficiency in listening and reading are not. All in all, the fit is quite good, especially since R^2 tends to be reduced with qualitative dependent variables.

Table 10
Model for Gain in Oral Proficiency (Preprogram OPI = 1+, 0/1 Criterion)

	OLS	
Variable	**Coefficient**	**t-Statistic**
Constant	186.719	3.76
Gender	0.24261	2.84
High school Russian	0.15325	1.72
Slavic languages	0.36988	1.99
Non-Slavic languages	0.09163	1.96
Program date	-0.09420	-3.77
% QualGen	0.00090	2.60

$R^2 = .252$, adjusted $R^2 = .212$, d.f. = 111.

3.6. *Summary: Factors Affecting Gains*

Table 11 summarizes the analysis of factors determining gains in all three modalities so that the effect of each variable across the board can be easily seen. While the statistical results are interesting, the most important findings from the point of view of second language acquisition can be summarized as follows:

— *Gender:* Men are more likely than women to gain on listening, and to go from 1+ to 2 and up on the OPI. See below in Section 4.2.
— *Age:* Younger students gain on listening.
— *Non-Slavic languages:* Knowing another foreign (non-Slavic) language is significant for OPI (and significant for listening and reading when MLATs

Table 11
Summary of Factors Affecting Gains on Three Language Modalities

Variable	ETS Listening	ETS Reading	OPI (0/1)	OPI (0/1/2)	OPI 1+ to 2↑
Gender	2.04	—	—	—	2.84
Age	-4.25	—	—	—	—
High school Russian	-1.68	—	2.03	2.23	1.72
College Russian	—	—	—	—	—
Slavic languages	-1.60*	—	—	—	1.99
Non-Slavic languages	*	*	—	1.48	1.96
Previous immersion	2.23	—	—	—	—
Major	—	—	—	2.13	—
Program date	—	—	-3.55	-3.93	—
MLAT3	4.45	*	—	—	—
MLAT4	*	*	—	—	—
MLAT5	—	—	—	—	—
MLATSF	—	6.73	—	—	—
QualGram	—	—	—	—	—
QualRead	*	3.40	—	—	—
QualGen	*	—	2.00	2.16	2.60
ETSL1	-16.50	1.66	—	—	—
ETSR1	7.80	-8.98	2.48	2.40	—
OPI1	—	—	-9.04	-9.72	—

Note: Numbers are t-statistics from "good" models.

* See Ginsberg 1992 for detailed data and analysis.

are not included);[21] the more non-Slavic languages known, the more the gain.

— *MLATs:* MLATs do not predict OPI. MLAT3 (analytic) and MLAT4 (synthetic) are good predictors of listening and reading. MLAT5 (memory) is not predictive. See below in Section 4.5.

— *Qualifying exams:* The grammar/reading qualifying exam as a whole predicts all modalities.22 See below in Section 4.4.

— *Preprogram level of modality:* On all modalities the higher the initial level, the less likely a gain; this is a function of the learning curve and the nature of the scale.

— *Preprogram reading proficiency:* Like the grammar/reading qualifying exam, higher preprogram reading proficiency seems to facilitate gains on the OPI and listening; the reading test may be picking up measurement errors, or the ceiling may be coming into play, but this is not likely.

— *Preprogram listening proficiency:* Preprogram listening proficiency does not seem to be very significant for the other criteria.

4. DISCUSSION

4.1. Second Language Acquisition Gains from Study Abroad

The traditional wisdom that one does not acquire real speaking competence without a period spent in-country is borne out by the ACTR data. Table 12 shows that only approximately 13 percent of students with four years of college Russian score an "advanced" level on the preprogram OPI, a figure that, curiously enough, does not vary greatly among students with two to five years of the language.

Assuming that the self-selected sample of Russian language students on which this study is based most likely represents the more successful of the students studying Russian in the United States, the maximum number of students with four years of college Russian reaching an advanced level in speaking is quite small indeed. Given the fact that the advanced level is considered only the minimal level of functional ability, it is clear that our educational system for the most part is not turning out students competent in speaking Russian.

By contrast, Table 5 shows that almost 40 percent of students return from the semester in Russia having reached a postprogram score of "advanced" in speaking. These data seem to indicate that at least one semester of study in-country is required if any sizable percentage of students studying Russian are to reach at least a functional level of competence in speaking. This judgment is not necessarily a criticism of the Russian language field, whether the teachers, students, or administrators. The language delivery system, so to speak, at the college level is not capable of dealing with a language of the degree of difficulty of Russian without significant supplementation, given the time of study it takes

to reach the "advanced" level. For example, if one takes the often quoted Foreign Service Institute scale concerning the amount of time required to learn Level 3 and 4 languages as a basis, assuming five hours a week of instruction in a college course, it would take eight years to equal the amount of instruction required by the FSI to produce a Level 2 speaker in Russian. This additional time could be added by extending the language-learning career down into secondary and primary school, by significantly intensifying instruction in college, or, most effectively, by building into the program a semester of language study abroad.

From the point of view of the student, this means that the overseas component must be planned for and built into the program at some point. Planning for this entails time and resource allocation above and beyond the normal college investment, and both students and parents must be made aware of this fact. From the point of view of teachers, program planners, and administrators, the obligatory inclusion of a study-abroad component requires close attention to articulation, both pre- and postprogram, as well as attention to scholarship resources for students with limited financial resources.

4.2. Gender

Perhaps the most controversial results of our analysis are those that clearly indicate that during their stay in Russia, women gain less than men in listening and speaking skills. In addition, men are more likely than women to cross the crucial divide between the intermediate-plus and advanced level.

The data supporting these conclusions are given in Tables 6 and 10. While the t-statistic in Table 6 is not overwhelming, it clearly indicates that gender is a significant factor in predicting success on the ETS Listening test. With regard to speaking ability, the results are more specific: Table 10 shows that men are more successful than women at passing through the advanced-level threshold, which is defined as genuine, although basic, functionality in the language. It need hardly be stated that results such as these must ordinarily be interpreted very carefully, particularly if the analysis used has few other factors built into it. However, given the range of factors controlled for in this study, such results raise a series of interesting questions and hypotheses that are perhaps overdue in the language-teaching profession.

Assuming that there are no real differences between men and women in language-learning ability—an assumption that everyone shares—three interpretations of these results come to mind:

— gender bias in the testing instruments
— a selection bias in the samples of men and women
— a difference in learning opportunities in-country

With regard to the first issue, the same instruments are used in both the pre- and postprogram tests, a fact that seems to weaken significantly the argument in

favor of a certain gender bias in these testing instruments. Presumably, any bias depressing the scores for women should be revealed in differences between men and women in initial preprogram scores. If, though, the bias were present only at the higher levels, it would account for depressed reflection of gain on the part of women in the postprogram scores, which naturally are at the higher-level range. Evidence for such speculation is provided precisely by the OPI scores as noted above, which show a difference between men and women at the highest level that most students attain: the intermediate-plus/advanced level. We stress, however, that our data only indicate a problem; an examination of possible gender bias in these OPI and ETS Listening testing instruments lies beyond the scope of this project. Related to, but independent of, the issue of gender bias is the question of the kinds of skills tested and accordingly reflected in the test scores. Not infrequently, for example, a semester in-country results in no increase in the OPI test score when the student is already at the advanced level or above coming into the program. (Of the eighty-five students who entered the program at an advanced level or higher, only thirty-one registered any gain on the OPI.) We attribute this fact to the nature of the learning curve, as mentioned above. OPI testers used in this study, though, have stated that students who test at the advanced level, for example, both at the beginning and at the end of the program often return as better speakers than they were when they left. Such evidence, albeit anecdotal, indicates that there are skills acquired that our testing instruments simply do not register—perhaps were not even designed to register. These unregistered skills may indeed be those that women acquire better than men, and that if included in the test results would negate the apparent advantage men seem to enjoy in the present situation. Unfortunately, we simply have no concrete evidence for this, although some effort is being made to examine this question in the Ford Foundation–sponsored part of this study.

A second possibility for explaining the gender difference is that self-selection is working to insure a greater number of more highly motivated or gifted males according to some qualities that are not reflected in the variables included in the study. This possibility for the time being remains unexplored.

Quite apart from artifacts in the tests, the gender differences might result from the fact that men and women have different learning experiences while abroad. For example, one might hypothesize that men and women spend their time differently in-country, and that therefore the difference in acquisition might be due to a difference in time-on-task. On the other hand, one might imagine a difference between the interactions of American men and women with Russian men and women as a reflection of the two cultural differences involved: American and Russian, as well as male and female. Finally, one might pose the question of whether the difference noted between men and women is a fact of study abroad in Russia, or whether it is observable—if the research were actually to be done—in learning situations in the United States as well as in other areas of the world where U.S. students go to study language. Indeed, the "chilly climate" is well documented in domestic classrooms, although the specific impact on language learn-

ing is not. Also, there is no reason to believe that Russia is the only society where our students study and where women receive differential treatment that might have an effect on language acquisition. While we are not prepared to discuss these issues here, a Ford Foundation–sponsored study of the ethnography of language learning in the study-abroad environment is currently being completed at the National Foreign Language Center.23

4.3. *Expert Language Learners*

The data in Tables 6, 9, and 10 show that students gain more in-country if they have had another foreign language in addition to Russian in high school or college. It seems logical that students would get better at learning foreign languages the more experience they have had at the task. This observation, however, does not tell us just what the advantage is. What are the language-learning skills that seem to come with experience in foreign languages? Do experienced language learners simply use more learning strategies, metacognitive, cognitive, and socioaffective? Or is it the case that they make more effective use of learning strategies, or stress certain ones at the expense of others? Do they have better communication strategies, which let them cope with difficulties and turn them into learning opportunities? Is it merely that their expectations are more realistic, and that frustrations are therefore attenuated and the rewards of perseverance anticipated? Once again, we must turn to ethnographic data for the answers.

From the policy point of view, the identification of advantages accruing to experienced language learners has important ramifications. First, it is worth noting that approximately 75 percent of the students in this study have studied another foreign language. Data from surveys of Chinese and Japanese conducted by the National Foreign Language Center confirm the tendency of students to come to these languages having had prior foreign language experience.[24] Accordingly, while the dearth of teachers and materials makes the inclusion of less commonly taught languages like Russian in this effort unrealistic—at least for the immediate future—these data argue for the particular advantage of any elementary school language study for even these languages if students are encouraged to try a second foreign language when they get to junior high school.

Finally, above the data showed that previous immersion also correlated significantly with gain in listening skills. This points to general experience in managing the virtual flood of input with which a student must cope in the in-country environment. Again, the question arises as to effective strategy selection and use (metacognitive, cognitive, socioaffective, communicative).

4.4. *Grammar and Reading Knowledge*

As previously reported (Brecht, Davidson, and Ginsberg 1990, 1991), qualifying grammar and reading achievement scores show significant predictive value for

speaking proficiency, reading proficiency, and listening proficiency alike. As noted in Table 11, higher levels of control of basic grammar and reading skills, as measured by QualGram/QualRead (replaced in 1990 with a single QualGen achievement test score in the same two modalities), are positively related to gains in all three skills. Of particular importance here is the strength of the relationships for gains in OPI across different levels and combinations of levels. While significant for all OPI gains, grammar/reading achievement proved to be most significant (t-statistic +2.6) for the group of learners at the 1+/2 speaking threshold, precisely the critical level in speaking proficiency for the greatest numbers of Americans studying Russian abroad.

The importance of this finding and the size of the sample on which it is based have obvious significance for the traditional debates concerning the role of explicit grammatical knowledge in the development of communicative skills (especially following Higgs and Clifford 1982). Most of the research in this debate has concerned the effect of instruction on accuracy and developmental sequences, on acquisition processes, on the rate of acquisition, and on the level of ultimate second language attainment.[25] Relatively little attention has been given to date to the practical concern that has beset classroom teachers and textbook authors alike for decades: What is the proper place of the study of explicit grammar, and if it is necessary, how much of what grammar? In asking these questions, teachers may not be concerned with how grammatically accurate their students will become in expressing themselves, as important as this is at the advanced levels. Rather, they simply want to know whether grammar instruction is worth the investment of time, especially of classroom time that might otherwise have been devoted to communicative skill practice.

The data in the current study provide the first empirical evidence that investment in grammar instruction in the early years of instruction may result in advances in speaking and listening skills at the upper-intermediate and advanced levels. This information is particularly important for teachers and textbook designers. In particular, if one views formal instruction as only one component in a student's language-learning career, formal instruction in grammar can be seen as one key element in producing expert language learners who will develop the independent capacity to gather and assimilate information and skills on their own through contact with native speakers. This is not to say, however, that we should undo the good that the focus on proficiency and functional competence has produced in the last decade or so. As the Interagency Language Roundtable–ACTFL guidelines indicate, a balance among functional skills, topical skills, and grammatical accuracy is the appropriate approach. These data simply underline the necessity to keep knowledge of grammar an equal partner in the goals of learning and instruction—not only for the skills they directly represent, but also for the good that accuracy does in advancing speaking and listening.

4.5. *MLATs*

The question of the predictive power of language aptitude, as registered in this project by MLAT3 (analytic), MLAT4 (synthetic), and MLAT5 (memory), has been the subject of much discussion ever since the appearance of Carrol and Sapon's Modern Language Aptitude Test in 1959 and Pimsleur's Language Aptitude Battery in 1966. Both "have been shown to make a consistent and substantial contribution to the prediction of student achievement in a variety of adolescent and adult language training programs" (Wesche 1981, p. 120, as quoted by Larsen-Freeman and Long 1991, p. 169). However, Larsen-Freeman and Long go on to say, "It is plausible that language aptitude tests work well to predict success because there is a concordance of tasks between the test and formal classroom study . . . , not because the test is measuring some innate linguistic ability" (p. 169). This assertion is captured, they continue, in the distinction by Cummins (1979) between cognitive/academic language proficiency (CALP) and basic interpersonal communication skills (BICS).

The results of the current study, we believe, provide important empirical support for this position. As is indicated in Table 11, MLAT3 (analytic) and MLAT4 (synthetic) are good predictors of listening and reading but do not predict speaking as reflected in the OPI. Accordingly, in the immersion experience where speaking is a primary goal, the MLAT is basically of no value in student selection or, for that matter, program design. (For this reason ACTR has ceased to administer the MLAT in its preprogram battery of tests.)

4.6. *Other Factors*

Space does not permit extensive further discussion of the other results summarized in Table 11. We shall, though, point out two more interesting issues raised by the data here, which are likely to be of interest in future research. One is the interrelationship of gains in listening and speaking skills:

— *High school Russian:* Having taken Russian in high school is negatively related to listening, positively related to OPI. This result may be an artifact of the correlations among other variables.
— *Slavic languages:* Knowing a Slavic language other than Russian is significant for OPI 1+ to 2 and above, but not significant for listening unless MLATs are included.
— *Previous immersion:* Previous immersion is positive for listening but not significant for OPI.

One might expect that these skills, being directly associated from the point of view of normal discourse, would be susceptible to the same predictors. Nonetheless, high school Russian, Slavic languages, and previous immersion, together

with age and MLATs, comprise a set of unrelated factors where the predictors are significantly different for these two skills.

Moreover, the threshold to minimal functional competence (1+ and 2) emerges as a focus for two important factors: gender and knowledge of another related (Slavic) language. That these two variables emerge especially strongly at this pivotal point in language ability seems unlikely to be coincidental.

CONCLUSION

In this paper we have presented a comprehensive analysis of the data collected by ACTR over a period of many years on students studying Russian in Moscow and St. Petersburg. The results displayed above are strong, often striking, but sometimes puzzling and subject to multiple interpretations. In any case, they constitute a corpus of empirically well grounded phenomena concerning adult language acquisition that need to be further explained, thereby setting the agenda for future research. For example, the fact that gender is significant only forces one to ask, Is it really? If so, why? What learning mechanisms are involved? How can the effect be mitigated? The answers to these questions require more study, including study of an ethnographic kind, with the purpose of seeing what actually happens in-country in the learning process.

NOTES

1. See DeKeyser 1991 and Freed 1990 and the references cited there, as well as Carlson et al. 1990 and Teichler and Opper 1988, for broader studies that include some discussion of language acquisition in the study-abroad environment.

2. This model of language acquisition in a foreign country contrasts with what is called "spontaneous second language acquisition," for which a major set of studies exists. See, for example, Perdue 1984.

3. From the beginning, the data collected by ACTR has not been limited to that needed for executing the academic exchanges; rather, academic policy ramifications were always a consideration. The present project was designed, first, to establish the basis for determining the success of a program with regard to quality and cost-effectiveness and to evaluate existing programs; second, to improve program design, implementation, and cost-effectiveness; and third, to determine the best predictors of success in a language-learning career, particularly at the more advanced levels, with the ultimate goal of selection and placement of students in specific programs. A further goal was added to the present U.S. Department of Education–funded project—that is, to establish and make available to different (particularly less commonly taught) languages a model for the collection of empirical data on second language acquisition.

4. The ACTR database, while possibly the most extensive of its kind in the field, does not contain the standard motivational measures. The academic program officers' ratings contain motivational assessments that may serve as proxies for these. See Brecht, Davidson, and Ginsberg 1990 and 1991 for discussion.

5. In fact, the current database contains over 900 cases, 658 of which are complete and in use here.

6. The resulting escalation of de facto qualifying standards worked to the advantage of students from institutions with intensive summer training and considerable advanced-level formal coursework (competency-based language courses beyond the level attained in the third year of college), effectively limiting to a significant degree the number of participants from smaller Russian departments throughout the United States.

7. Students over age thirty-five, who know more than one Slavic language besides Russian, who have had more than one previous immersion experience, who studied in Russia in the previous semester and are hence in effect in the second half of a ten-month program, or who have no change on any measure of the three criteria used below—such students have been dropped from the analysis. Data are available on an additional 182 students attending programs in the fall of 1990 and spring of 1991 but are not complete enough to include in this analysis.

8. Means, simple cross-tabulations, scatterplots, and so on, which are not controlled, must, however, be treated as specific to the sample at hand. By way of orientation to the analysis, it should also be pointed out that a key factor

determining language gains—namely, what the students do while in Russia—is not part of the ACTR database, so we will not address here why the results hold. Nevertheless, hold they do. As noted above, in subsequent research based on the National Foreign Language Center/Ford Study Abroad Project we shall attempt to account for some of the patterns presented below.

9. These instruments were selected because they are widely accepted as standard, thereby offering the advantage of general accessibility and comparability. Whether they meet sufficient validity and reliability criteria for measuring Russian language skills is a separate question, which is touched upon in 4.2.

10. The combination is weighted by the number of questions on the reading and grammar parts and is accordingly dominated by the grammar part.

11. As discussed in Brecht, Davidson, and Ginsberg (1990, 1991), ACTR resident directors at each institute were asked to rate students on a scale of 1 (lowest) to 5 (highest) on seven criteria, comparable to the kinds of ratings one finds on a recommendation to graduate school: intellectual motivation, natural ability to learn, willingness to use Russian, cultural adaptability, taking advantage of cultural opportunities, ability to work in a group, and leadership potential. For reasons of space, the effects of these variables—the only measures of motivation and attitudes available—are explored in a separate paper. See Ginsberg 1992, pp. 20–22.

12. Preliminary studies showed that the imputation does not affect the results. Of course, missing data on the language measures (dependent and key independent variables) never were imputed, and these cases were accordingly dropped from the relevant analyses.

13. As noted, regressions of gains in ETS Listening and Reading are estimated by OLS. The methods of choice for estimating effects of gains on the OPI, a qualitative variable, are probit and ordinal probit regression, depending on whether the criterion is 0/1 (no gain/gain) or 0/1/2 (no gain/gain of one step/gain of two or more steps). Logit and ordinal logit models could also be used, but they yield almost exactly the same results as the probits. Indeed, as will be seen in Tables 6 and 7, OLS— which is used to get initial values for the probit estimations—gives essentially the same levels of significance as the probits (and logits) themselves. (Coefficients are not directly comparable because of normalizations in the probits, but the ratios of coefficients can be compared.) That is, probits, logits, and OLS lead to the same qualitative conclusions. Accordingly, when variables are being screened for effects, we simply report the OLS results—which in any event are easier for many readers to interpret. For the good models, both OLS and probit (ordinal probit) coefficients and t-statistics are presented.

14. Much of the regression analysis was done in Data Desk, although some of the more complex models were computed using Crawtran and BMDP. A complete set of tables is available in Ginsberg 1992.

15. For the ACTR data there is reason to consider ETSL2, rather than the gain, as a criterion because the two ETS tests of listening proficiency are not of the same difficulty (see Brecht, Davidson, and Ginsberg 1991, paragraph 1.2 and

note 5). As a technical matter, in OLS regression it does not matter whether the dependent variable is taken as gain itself or the postprogram score, as both yield the same results—that is, the same residual sums of squares, coefficients, and t-statistics—which in this section are our primary concern. More complex latent variable models, which are not warranted here, would use ETSL2 as dependent. Gain is fine for ETSR, for which the two tests are equivalent.

16. Allowing for measurement error in the measured OPI score itself introduces another level of complexity that is not considered here.

17. A four-level criterion (zero, one, two, three, or more half-points) was also explored but turned out to be equivalent to the three-level criterion. An adjustment of the raw half-points gained, reflecting the difficulty of gaining as a function of the preprogram level (see Brecht, Davidson, and Ginsberg 1990, 1991), was also explored in some detail but found to be essentially equivalent to the 0/1 and 0/1/2 criteria.

18. See below in Sections 3.3 and 3.4, concerning the relevance of previous foreign language experience to listening and reading skills.

19. See Ginsberg 1992, Figure 6.

20. Against this strategy is the fact that the numbers in each group are perforce smaller than the whole, and that it is difficult to estimate common structure. With the ACTR data, however, the numbers in the subgroups are still much larger than those used in other studies in the literature.

21. See Ginsberg 1992, pp. 15–18.

22. The grammar part by itself does not add anything to reading.

23. This project employs self-report instruments in the form of diaries and written and oral journals, collected from students spending a semester in-country studying Russian. Participant and nonparticipant observation and focused and nonfocused interviews are being analyzed for evidence of differences in input and intake, as well as caretaking on the part of Russian men and women. These data are providing some indication of the issues involved, which will be reported in a series of papers to be released by the National Foreign Language Center.

24. Cf. Moore with Lambert and Walton 1992; and Jorden with Lambert 1991.

25. See Larsen-Freeman and Long 1991, pp. 299ff.

REFERENCES

Brecht, R., and D. Davidson. 1992. "Language Acquisition Gains in Study Abroad: Assessment and Feedback." In *Language Assessment for Feedback: Testing and Other Strategies*, ed. E. Shohamy and A. R. Walton, pp. 87–101. Dubuque, Iowa, and Washington, D.C.: Kendall/Hunt and the National Foreign Language Center.

Brecht, R.; D. Davidson; and R. Ginsberg. 1990. "The Empirical Study of Proficiency Gain in Study Abroad Environments of American Students of Russian." In *American Contributions to the VII International Congress of MAPRIAL*, ed. D. Davidson, pp. 123–52. Washington, D.C.: American Council of Teachers of Russian.

——. 1991. "On Evaluating Language Proficiency Gain in Study Abroad Environments: An Empirical Study of American Students of Russian (A Preliminary Analysis of Data)." In *Selected Papers Delivered at the NEH Symposium in Russian Language and Culture*, ed. Z. D. Dabars, pp. 101–30. Baltimore: CORLAC, Friends School of Baltimore.

Carlson, J. S., et al. 1990. *Study Abroad: The Experience of American Undergraduates*. New York: Greenwood.

Cummins, J. 1979. "Cognitive Academic Language Proficiency, Linguistic Interdependence, the Optimum Age Question, and Some Other Matters." *Working Papers on Bilingualism* 19:197–205.

DeKeyser, R. 1991. "Foreign Language Development during a Semester Abroad." In *Foreign Language Acquisition Research in the Classroom*, ed. B. F. Freed, pp. 104–119. Lexington, Mass.: D. C. Heath and Co.

Educational Testing Service, International Testing and Training Program Division. 1984. "Advanced Russian Listening/Reading Proficiency Test." Princeton: Educational Testing Service.

Freed, B. 1990. "Language Learning in a Study Abroad Context: The Effects of Interactive and Non-interactive Out-of-Class Contact on Grammatical Achievement and Oral Proficiency." In *Linguistics, Language Teaching, and Language Acquisition: The Interdependence of Theory, Practice, and Research*, ed. J. E. Alatis, pp. 459–477. Georgetown University Round Table on Languages and Linguistics 1990. Washington, D.C.: Georgetown University Press.

Ginsberg, R. 1992. *Language Gains during Study Abroad: An Analysis of the ACTR Data*. NFLC Working Papers. Washington, D.C.: National Foreign Language Center.

Higgs, T., and Clifford, R. 1982. "The Push toward Communication." In *Curriculum, Competence, and the Foreign Language Teacher*, ed. T. Higgs, pp. 57–79. Skokie, Ill.: National Textbook Company.

Jorden, E. H., with R. D. Lambert. 1991. *Japanese Language Instruction in the United States: Resources, Practice, and Investment Strategy*. NFLC Monograph Series. Washington, D.C.: National Foreign Language Center.

Larsen-Freeman, D., and M. H. Long. 1991. *An Introduction to Second Language Acquisition Research.* New York: Longman.

Moore, S. J., with R. D. Lambert and A. R. Walton. 1992. *Introducing Chinese into High Schools: The Dodge Initiative.* NFLC Monograph Series. Washington, D.C.: National Foreign Language Center.

Perdue, C., ed. 1984. *Second Language Acquisition by Adult Immigrants: A Field Manual.* Rowley, Mass.: Newbury House.

Teichler, U., and S. Opper. 1988. *Erträge des Auslandsstudiums für Studierende und Absolventen.* Bonn, Germany.

Wesche, M. 1981. "Language Aptitude Measures in Streaming, Matching Students with Methods, and Diagnosis of Learning Problems." In *Differences and Universals in Language Learning Aptitude,* ed. K. Diller, pp. 119–54. Rowley, Mass.: Newbury House.

ABOUT THE AUTHORS

Richard D. Brecht received his M.A. and Ph.D. in Slavic languages and literatures from Harvard University and is currently professor of Slavic languages and literatures at the University of Maryland at College Park. Co-founder of Project ICONS (International Communication and Negotiation Simulations), he is a senior research associate at the National Foreign Language Center, director of research and development at the American Council of Teachers of Russian/American Council for Collaboration in Education and Language Study, and co-director of the National Council of Organizations of Less Commonly Taught Languages. He has authored numerous books and articles on Slavic linguistics, Russian language, and second language acquisition.

Dan Davidson is professor of Russian and second language acquisition at Bryn Mawr College and executive director of the American Council of Teachers of Russian/American Council for Collaboration in Education and Language Study in Washington, D.C. He received his M.A. and Ph.D. in Slavic languages and literatures from Harvard University and has been involved in U.S.-Soviet academic cooperation as a teacher, writer, and academic exchange specialist for the past twenty years. He is the author of several books and studies in the area of Russian and second language acquisition, including a fifteen-year longitudinal study of skill gains by American students enrolled in Russian language training programs.

Ralph B. Ginsberg is professor of regional science at the University of Pennsylvania, where he was chairman of the department for eight years. He has been a visiting professor at Erasmus University in the Netherlands, Umeå University in Sweden, and the Woodrow Wilson School at Princeton University. He is a senior research associate at the National Foreign Language Center and has been a deputy director of the center. He has written extensively on probability models and data analysis, and is the author of several large-scale empirical studies. He received his Ph.D. in sociology at Columbia University and edited the *Journal of Mathematical Sociology*. He has served on the boards of several major journals. His current research concerns language learning in computer-based multimedia settings and language study abroad.

The National Foreign Language Center

at the Johns Hopkins University
1619 Massachusetts Avenue NW
Washington DC 20036

BEST COPY AVAILABLE